

확률강우량의 공간분포 추정을 위한 KED 기법의 적용

Application of KED Method for Estimation of Spatial Distribution of Probability Rainfall

서 영 민* / 여 운 기** / 이 승 윤*** / 지 홍 기****

Seo, Young Min / Yeo, Woon Ki / Lee, Seung Yoon / Jee, Hong Kee

Abstract

This study employs the KED method using the correlations between probability rainfall and topographical factors as single auxiliary variable for assessing the effectiveness of external variables to improve the reliability in the estimation of spatial distribution of probability rainfall. As a result, the KED method gives similar results compared with deterministic spatial interpolation methods and kriging methods in the estimation of rainfall spatial distribution and mean areal rainfall, and as a result of the cross-validations of KED and kriging methods, the KED method using terrain elevation as auxiliary variable gives the best results, which are not significantly different in comparisons with other methods.

Keywords : probability rainfall, spatial distribution, topographic factor, KED, kriging

요 지

본 연구는 확률강우량에 대한 공간분포 추정시 신뢰도를 향상시키는데 있어서 외부변수 사용의 유효성을 평가하기 위하여 확률강우량과 단일 보조변수로서 지형특성인자들과의 상관관계를 고려한 KED 기법을 적용하였으며, 그 결과 강우공간분포 및 유역평균강우량의 추정에 있어서 확정론적 공간보간기법 및 크리징 기법과 대체로 비슷한 결과를 나타내는 것으로 분석되었으며, KED 및 크리징 기법에 대한 교차검증 결과 보조변수로서 표고를 사용한 KED 기법이 가장 좋은 결과를 나타내고는 있으나 다른 기법들과 비교했을 때 큰 차이를 보이지 않는 것으로 분석되었다.

핵심용어 : 확률강우량, 공간분포, 지형특성인자, KED, 크리징

1. 서 론

강우의 공간분포에 대한 신뢰성 있는 추정은 수자원

해석 및 설계에 있어서 필수적인 요소이다. 강우장의 공간 변동성에 대한 고해상도 추정은 홍수, 특히 돌발홍수의 원인이 되는 국지성 호우의 확인 및 분석에 있어서 특

* 교신저자, 영남대학교 건설시스템공학과 박사과정 (e-mail: elofy@nate.com)

Ph.D., Student, Department of Civil Engineering, Yeungnam Univ., Gyeongsangbuk-do 712-749, Korea

** 영남대학교 건설시스템공학과 박사과정 (e-mail: adonas@nafree.net)

Ph.D., Student, Department of Civil Engineering, Yeungnam Univ., Gyeongsangbuk-do 712-749, Korea

*** 한국수자원공사, K-water 수자원연구원 선임연구원 (e-mail: leesy@kwater.or.kr)

Senior Researcher, K-water Institute, 462-1 Jeonmin-dong, Daejeon 305-730, Korea

**** 영남대학교 건설시스템공학과 교수 (e-mail: hkjee@yu.ac.kr)

Professor, Department of Civil Engineering, Yeungnam Univ., Gyeongsangbuk-do 712-749, Korea

히 중요하다. 또한 강우의 공간 변동성에 대한 고려는 면적평균강우량 추정의 정확도를 향상시키는데 있어서 중요하다. 강우-유출모델의 모의결과에 대한 신뢰도를 향상시키는데 큰 영향을 미친다.

우리나라는 국토의 67%가 산지로 이루어져 있기 때문에 산악 등 지형에 의한 국지기상변화가 매우 심하며, 여름철 산악지형에서 발생하는 국지성 집중호우나 태풍에 의한 집중호우가 좋은 예이다. 산악지형의 경우 접근성 및 강우측정의 어려움, 강우공간분포의 복잡성 때문에 강우관측망이 조밀하지 않을 경우 산악지형에 대한 국지적 강우 변동성을 규명하기에는 한계가 있기 마련이다. 강우의 공간 변동성을 특성화하기 위해서는 강우관측망을 균등하게 분포시키고 그 밀도를 증가시키는 것이 가장 좋은 해결방법이나 방대한 설비비용 및 비용문제 때문에 현실적으로는 매우 어려운 것이 현실이다. 그리고 강우공간분포 측면에서 최근 레이더 자료를 이용하여 강우량을 추정하는 연구가 활발히 진행되고 있으나 복잡한 산악지형으로 이루어진 지역에 있어서 정량적 강우정보를 정확하게 추정할 수 있을 정도로 레이더 자료의 공간적 해상도가 정밀하지 않으며, 강우-유출모델의 모의에 있어서 사용되는 다른 GIS 정보와의 공간적 해상도 차이 및 우량계 자료와의 결합에 따른 유출량 모의에 있어서의 불확실성 문제가 여전히 여러 연구들에서 제기되고 있는 실정이다 (Cole and Moore, 2008; Moulin et al., 2009).

따라서 강우의 공간분포에 대한 대부분의 연구들은 강우관측지점의 관측치를 이용하여 미계측 지점의 값을 공간적으로 예측하는 공간보간기법들에 집중되고 있다. 강우의 공간분포를 추정하기 위한 전통적인 공간보간기법으로는 Thiessen 다각형 (또는 Voronoi 다각형), IDW (inverse distance weighted) 및 등우선법 등을 들 수 있다. 공간보간기법은 크게 확정론적 방법과 지구통계학적 방법 (또는 공간통계학적 방법)으로 구분할 수 있다. 확정론적 방법으로는 Thiessen 다각형, IDW, 경향면 분석 (trend surface analysis), Spline 등이 있으며, 지구통계학적 방법으로는 크리깅 (kriging) 기법, 즉 단순크리깅 (simple kriging), 보통크리깅 (ordinary kriging), 일반크리깅 (universal kriging), KED (kriging with external drift), 회귀크리깅 (regression kriging) 등이 있다. 공간보간기법에 대한 기타 분류방법은 Hartkamp et al. (1999)을 참고할 수 있다.

강우는 시간 및 공간에 따라 변동하며, 이러한 변동성은 시간 스케일이 작을수록 증가한다. 따라서 작은 시간 스케일에 대하여 제한된 수의 관측치로부터의 공간보간은 용이하지 않다. 즉, 월 또는 연평균 강우량에 비하여 일평균강우량에 대한 공간보간은 증가된 변동성으로 인하

여 어려움이 따른다 (Haylock et al., 2008; Yatagai et al., 2009). 그리고 강우관측망은 일반적으로 조밀하지 않고 고르게 분포되어 있지 않으며, 대체로 계곡 및 표고가 낮은 평지에 위치한다. 즉, 표고가 낮은 평지에서의 강우관측망의 밀도가 높고 표고가 높은 지역에서는 강우관측망의 밀도가 낮은 것이 일반적이다. 가용한 강우관측자료는 낮은 표고에 대하여 편향되어 있으며, 이러한 이유로 강우패턴에 대한 모델링은 더욱 어려운 실정이다.

최근 많은 연구들에서 이러한 편향문제를 보완하기 위하여 보조변수를 도입하여 강우패턴 모델링에서의 정확도를 향상시키려는 시도가 이루어지고 있다 (Kieffer and Bois, 2000; Bostan and Akyurek, 2007). 예를 들어, 기후, 지리 및 지형인자와 강우량과의 상관관계를 회귀분석을 이용하여 분석하고 그 결과를 강우패턴 모델링과 결합하여 제한되고 편향된 강우자료를 이용함으로써 발생하는 모델링에서의 한계를 보완하는 것이다. 여기서 장기 누적시간에 대한 강우량, 즉 월평균 또는 연평균 강우량은 표고가 증가함에 따라 증가하는 것으로 알려져 있으나 단기 누적시간에 대한 강우량과 표고와의 상관관계는 일반적으로 낮은 것으로 알려져 있다 (Daly, 2002). 이러한 문제 때문에 기존 연구들에서는 강우량과 단일 지형인자간의 상관관계를 이용하는 것이 아니라 여러 지형인자들과의 중회귀분석을 적용하는 경우가 많다 (Buytaert et al., 2006; Krähenmann and Ahrens, 2010).

따라서 본 논문에서는 확률강우량에 대한 강우공간분포 및 유역평균강우량의 추정시 신뢰도를 향상시키는데 있어서 외부변수 사용의 유효성을 평가하기 위하여 강우공간분포 추정시 보조변수를 크리깅 시스템에 통합할 수 있는 KED 기법을 적용하였다. 그리고 확정론적 공간보간기법인 Thiessen 다각형, IDW, 경향면 분석 및 Spline과 지구통계학적 공간보간기법인 단순크리깅, 보통크리깅 및 일반크리깅과의 비교분석을 통하여 그 적용성을 평가하였다.

2. 공간보간기법

2.1 확정론적 공간보간기법

본 연구에서는 지구통계학적 기법과의 비교를 위하여 확정론적 공간보간기법인 경향면 분석, IDW, Thiessen 다각형 및 Spline 기법을 적용하였다.

경향면 분석은 측정지점의 좌표값을 이용하여 다중회귀모델을 구성한 후 구성된 다중회귀모델을 이용하여 미관측 지점의 값을 예측하는 공간보간 방법이며, IDW는 공간적으로 인접한 지점 사이의 값은 공통된 위치요인으로 인하여 유사성을 갖게 되는 반면에 두 지점 사이의 거

리가 증가할수록 이러한 유사성은 상대적으로 감소하게 된다는 가정에 기초한 기법으로서 거리가 증가함에 따라 그 값이 감소하는 가중치를 공간예측에 도입하며, 거리에 따른 가중치의 변화정도는 거리에 대하여 적용된 지수값에 의존한다. 그리고 Voronoi 다각형 또는 Dirichlet 격자라고도 불리는 Thiessen 다각형은 속성값이 알려진 이차원의 공간에서 특정 지점과 가장 인접한 지점의 속성값을 이용하여 특정지점에 대한 속성값을 추정하는 기법이며, Spline 보간법은 3차원 표면에 대하여 전반적인 표면곡률을 최소화시켜 입력 자료점을 정확하게 통과하는 매끄러운 표면을 생성시키는 수학적 함수를 이용하여 공간예측을 수행하는 확률론적 공간보간기법으로서 1차, 2차 및 3차 도함수를 최소연산으로 조합하는 Regularized Spline과 1차 및 2차 도함수만을 사용하지만 연산이 더 많은 점을 사용하기 때문에 부드러운 표면을 생성하지만 연산시간이 긴 Tension Spline으로 구분할 수 있다.

2.2 지구통계학적 공간보간기법

지구통계학적 공간예측은 공간중속성을 이용하여 미지의 위치에서의 값을 예측하는 기법으로서 지구통계학에서 널리 사용되는 크리깅은 예측값과 그와 관련된 분산 또는 표준오차를 추정할 수 있으며, 공간과정 (spatial process)을 특성화시키는 공간중속성에 대한 모델의 형태와 매개변수 추정을 필요로 한다. 공간중속성에 대한 모델은 관심지역 상에 있는 임의의 두 지점 사이의 거리에 대하여 표현되며, 이러한 모델들은 공분산 (covariance) 또는 베리오그램 (variogram)의 형태로 얻어진다.

이와 같은 공간예측은 다음의 두 단계로 수행된다. 첫 번째 단계는 공간과정의 공분산과 베리오그램을 모델링한다. 본 단계에서 경험베리오그램 (empirical variogram)을 계산하고 계산된 경험적 모델을 바탕으로 이론베리오그램 (theoretical variogram)의 형태와 관련 매개변수들을 결정한다. 그리고 두 번째 단계에서는 첫 번째 단계에서 얻어진 모델과 매개변수로부터 관심영역 상에서 크리깅 시스템을 통하여 예측값과 그와 관련된 분산 또는 표준오차를 추정하게 된다.

2.2.1 크리깅

크리깅은 임의의 확률변수에 대해서 측정된 위치에서의 측정값을 이용하여 측정되지 않은 위치에서의 확률변수에 대한 수치를 가중선형회귀 (weighted linear regression)로 예측하는 공간예측기법으로서 예측오차를 최소화하면서 가중치를 구하는 방법으로 수행되며, 이를 수식으로 나타내면 Eq. (1)과 같다.

$$\hat{z}(s_0) = \sum_{i=1}^n w_i z(s_i) \quad i = 1, 2, \dots, n \quad (1)$$

여기서, $\hat{z}(s_0)$ 는 측정되지 않은 위치 s_0 에서의 예측값을 의미하고 $z(s_i)$ 는 임의의 위치 s_i 에서 측정된 값들을 의미한다. 그리고 w_i 는 위치 s_i 에서의 가중치이다.

공간자료가 확률과정의 실현값이라고 할 때 확률벡터 $z(s)$ 는 Eq. (2)와 같이 모델링할 수 있다.

$$z(s) = m(s) + e(s) \quad (2)$$

여기서, $m(s)$ 는 방향 또는 추세를 나타내는 항이고 $e(s)$ 는 오차항을 의미한다. $m(s)$ 가 일정한 기지값일 경우 단순크리깅, 일정한 미지값일 경우 보통크리깅이라고 하며, $m(s)$ 가 공간추세를 가질 경우 일반크리깅이라고 한다.

한편 일반크리깅, KED 및 회귀크리깅은 하이브리드 크리깅 (hybrid kriging) 계열에 속하는 비정상 지구통계학적 기법 (non-stationary geostatistical method)으로서 기본적으로는 동일한 기법이다. 일반크리깅은 Matheron (1969)에 의해 처음으로 소개되었으며, 추세에 대한 평균이 일정하지 않은 크리깅의 특별한 경우로서 추세가 좌표만의 함수로서 모델링된다. 반면에 변동에 대한 확률론적 부분, 즉 추세항 (drift)이 좌표가 아니라 어떤 보조변수에 대한 선형함수로서 보조적으로 정의될 경우 이러한 크리깅을 KED라고 한다 (Wackernagel, 2003). 일반크리깅 또는 KED의 경우 잔차에 대한 공분산행렬이 보조예측변수에 의해 확장됨으로써 관심변수에 대한 예측치를 추정한다 (Webster and Oliver, 2001). 그러나 추세항 및 잔차는 분리하여 추정할 후 나중에 합산될 수 있으며, 이러한 절차는 Ahmed and de Marsily (1987) 및 Odeh et al. (1995)에 의해 제안되었고 이후에 회귀크리깅으로 명명되었다.

2.2.2 KED

KED는 대상지역 내에 적은 개수의 관측치만을 가지는 확률변수 Z 에 대하여 동일한 지역에 광범위하게 알려져 있는 다른 변수 s 를 이용하여 미지의 위치에서의 값을 예측하는 기법이다.

이러한 두 변량이 선형관계를 가진다고 가정하면, Eq. (3)과 같이 나타낼 수 있다.

$$E[Z(x)] = b_0 + b_1 s(x) \quad (3)$$

여기서, b_0 는 상수이고 b_1 는 계수이다.

2차 정상확률함수 $Z(x)$ 를 고려할 때 형상함수 (shape function) $s(x)$ 를 도입함으로써 $Z(x)$ 에 대한 평균표본간격보다 더 작은 스케일로 설명될 경우 그것의 예측치는 향상될 수 있다. 예측치 $z^*(x_0)$ 는 가중치 w_i 의 합이 1이

되는 위치 $x_i, i = 1, \dots, n$ 에서 표본값 $z(x_i)$ 에 대한 선형 조합이 되며, Eq. (4)와 같이 나타낼 수 있다.

$$z^*(x_0) = \sum_{i=1}^n w_i z(x_i), \quad \sum_{i=1}^n w_i = 1 \quad (4)$$

Eq. (5)는 일반크리깅 방정식 (universal kriging equation)이라고 불리는 연립선형방정식이며, Eq. (6)과 같은 최소예측분산 σ_K^2 을 가진다.

$$\begin{cases} \sum_{j=1}^n w_j C(x_i - x_j) - \mu_1 - \mu_2 s(x_i) = C(x_i - x_0) \\ \text{for } i = 1, 2, \dots, n \\ \sum_{i=1}^n w_i = 1, \quad \sum_{i=1}^n w_i s(x_i) = s(x_0) \end{cases} \quad (5)$$

$$\sigma_K^2 = C(0) - \sum_{i=1}^n w_i C(x_i - x_0) + \mu_1 + \mu_2 s(x_0) \quad (6)$$

여기서, C 는 공분산함수, $C(0)$ 는 거리 0에서의 공분산, μ_1 및 μ_2 는 라그랑지 매개변수이다.

KED는 크리깅 시스템의 공간 정의역에서 광범위하게 측정된 하나 또는 그 이상의 보조추세변수 (external drift variable) $s(x_i), i = 1, \dots, n$ 에 대한 추가적인 일반성 조건을 통합하는 것으로 구성된다. KED에서는 관측지점에서 관심변수와 보조변수 사이의 선형관계를 가정하며, 만

약 두 변수 사이의 관계가 비선형함수에 의해 설명된다면 먼저 보조변수에 대한 자료를 선형으로 변환한 후 보조추세로 사용될 수 있다 (Bourennane et al., 2000).

3. 결 과

본 연구는 KED 기법의 적용에 있어서 보조변수로서 지형특성인자를 고려한 확률강우량의 공간분포 및 유역 평균강우량 추정을 목적으로 한다. 이를 위하여 국제수문 개발계획 (IHP)에 의해 선정 및 운영된 바 있는 위천대표 유역을 포함하고 있어 양질의 수문자료를 보유하고 있으며, 유역의 대부분이 산지로 이루어져 있고 도시화율이 낮은 위천유역을 분석유역으로 선정하였다.

먼저 확률강우량을 추정하기 위하여 국가수자원관리종합정보시스템 (WAMIS)에서 제공하는 국토해양부 관할의 위천유역 인근 18개 강우관측소 (Table 1 및 Fig. 1 참조)에 대한 시우량 자료를 수집하였다. 지점별로 사용가능한 자료의 개수가 상이하고 결측 및 이상치가 포함되어 있기 때문에 RDS (Reciprocal Distance Squared) 기법을 이용하여 자료의 보완 및 확충을 실시하였으며, 이로부터 1980~2008년 (29년) 동안의 연최대강우량 계열을 작성하였다. 강우빈도분석을 위해 FARD 2006을 이용하여 연최대강우량 계열에 대한 예비적 해석 (기본통계치 산정, 무작위성 검사), 확률분포형 별 매개변수 산정 (모멘트법, 최

Table 1. Topographic Factors and Probability Rainfall at Raingauges

No.	Stations	Easting	Northing	Elevation (EL.m)	Aspect	Hillshade	Slope (%)	R100 (mm)
1	Angae	329,162	321,316	59	210.35	180	0.59	313
2	Daeyul	349,087	285,931	216	34.44	181	5.07	320
3	Doriwon	341,826	312,202	60	257.21	180	0.53	280
4	Gomae	342,443	289,674	151	303.46	181	0.46	421
5	Hwasan	359,404	289,936	624	227.51	168	42.61	294
6	Hwasu	357,634	293,727	160	265.47	180	2.07	311
7	Hyeonseo	371,888	307,903	525	201.14	154	27.92	330
8	Hyoryeong	342,785	296,278	103	336.43	185	3.24	324
9	Jangchun	334,556	293,249	54	240.90	185	15.87	276
10	Kunwi	341,330	305,843	108	212.46	179	2.18	293
11	Kunwi Nambu	340,661	300,868	103	308.15	184	2.39	321
12	Sanseong	352,690	293,239	140	264.79	189	8.93	344
13	Seobu	355,864	286,388	181	107.72	177	1.58	277
14	Seoksan	364,337	297,767	222	251.11	181	1.60	302
15	Shinryeong	361,185	284,264	139	211.73	179	0.85	232
16	Uihung	354,222	298,569	139	166.12	180	0.03	345
17	Uiseong	351,548	318,595	98	304.39	220	27.33	314
18	Woobo	349,287	300,642	120	319.44	181	0.49	323

※ R100: Probability rainfall for return period 100 yr with duration 24hr, Light's position for hillshade: Azimuth=315°, Altitude=45°



Fig. 1. Study Area and Location of Raingauges

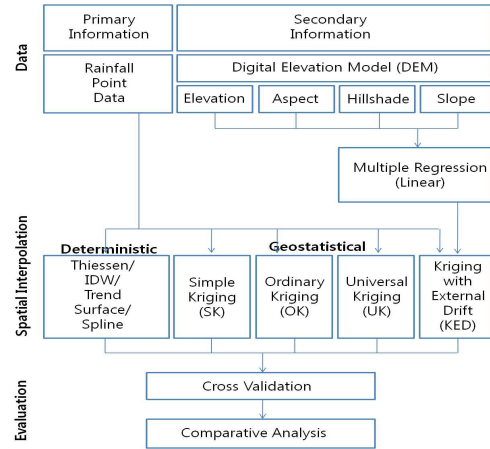


Fig. 2. Flow Chart for Rainfall Spatial Interpolation

Table 2. Summary Statistics of Probability Rainfall and Topographic Factors at Raingauges

Statistics	R100 (mm)	Elevation (EL.m)	Aspect	Hillshade	Slope (%)
MIN	232.1	54.0	34.4	154.0	0.03
MEDIAN	313.5	139.0	246.0	180.0	2.12
MEAN	312.2	177.9	234.6	181.3	7.98
MAX	420.8	624.0	336.4	220.0	42.61
SD	38.6	152.9	76.3	12.2	12.33

※ MIN: Minimum, MAX: Maximum, SD: Standard Deviation

우도법, 확률가중모멘트법), 매개변수의 적합성 검토, 확률분포형의 적합도 검정 (χ^2 검정, Kolmogorov-Smirnov 검정, 확률도시 상관계수검정, Cramer Von Mises 검정), 확률강우량 산정과 같은 절차를 수행하였으며, 그 결과 최적확률분포형으로 Gumbel 분포를 채택하였다.

지형특성인자(표고, 향, 음영기복, 경사)에 대한 변환, 탐색적 자료분석 및 공간보간분석을 실시하기 위하여 ESRI에서 개발된 GIS 분석 소프트웨어인 ArcView 및 ArcGIS와 통계계산 및 그래픽을 목적으로 GNU 프로젝트로 개발된 프로그래밍 언어인 R을 이용하였다. 지형특성인자에 대한 자료는 국가수자원관리종합정보시스템에서 제공하는 지형공간 래스터 파일 (30×30m)을 이용하였으며, 각 강우관측소에 대한 지형특성인자 및 100년 빈도 확률강우량(지속시간 24hr)은 Table 1과 같다.

본 연구에서는 강우공간분포 및 유역평균강우량을 추정하기 위하여 지구통계학적 기법, 즉 단순크리깅, 보통크리깅, 일반크리깅 및 KED 기법을 적용하였다. 그리고 지구통계학적 기법과의 비교를 위하여 확정론적 보간기법인 Thiessen 다각형, IDW, 경향면 분석 및 Spline 기법을 적용하였다. 본 연구에서의 분석절차에 대한 개요는 Fig. 2와 같다.

3.1 탐색적 자료분석

Tables 2 and 3은 강우관측지점에서의 100년 빈도 확률강우량(지속시간 24hr) 및 지형인자들의 요약통계치와 유역전체에 대한 지형인자들의 요약통계치를 분석한 결과이다. 강우관측지점의 경우 표고가 EL. 54.0~624.0m, 향이 34.4~336.4, 음영기복이 154.0~220.0, 경사가 0.03~42.61%의 범위를 나타내는 반면 유역전체의 경우 표고가 EL. 37.0~1,186.0m, 향이 -1.0~360.0, 음영기복이 0.0~254.0, 경사가 0~158.5%의 범위를 나타내는 것으로 분석되었다.

Table 4는 강우관측지점에서 100년 빈도 확률강우량 및 지형특성인자들 사이의 상관계수를 산정한 것으로서 확률강우량과 지형인자들 간의 상관계수를 살펴보면 상관계수의 값이 0.25 이하로서 상관관계가 매우 약한 것으로 분석되었다. 그리고 지형인자들 간의 상관계수를 살펴보면 X~EL, EL~HLS, EL~SLP가 각각 0.67, -0.60, 0.74로서 중속성을 보이는 것으로 분석되었다.

3.2 중회귀분석

강우의 공간분포를 추정하는데 있어서 기존의 강우관측망은 신뢰성 있는 결과를 산출하기 위해 충분히 조밀하

Table 3. Summary Statistics of Topographic Factors in Wichun Watershed

Statistics	Elevation(EL.m)	Aspect	Hillshade	Slope(%)
MIN	37.0	-1.0	0.0	0.0
MEDIAN	178.0	184.6	179.0	28.2
MEAN	210.0	180.9	171.9	29.1
MAX	1,186.0	360.0	254.0	158.5
SD	134.7	106.4	40.4	19.6

Table 4. Correlation Matrix for Topographic Factors at Raingauges

Variables	R100	X	Y	EL	ASP	HLS	SLP
R100	1.000	-0.117	0.049	0.028	0.254	0.046	-0.067
X	-0.117	1.000	-0.314	0.665	-0.200	-0.335	0.350
Y	0.049	-0.314	1.000	-0.211	0.304	0.268	0.095
EL	0.028	0.665	-0.211	1.000	-0.225	-0.600	0.738
ASP	0.254	-0.200	0.304	-0.225	1.000	0.331	0.026
HLS	0.046	-0.335	0.268	-0.600	0.331	1.000	-0.059
SLP	-0.067	0.350	0.095	0.738	0.026	-0.059	1.000

※ X: Easting, Y: Northing, EL: Elevation, ASP: Aspect, HLS: Hillshade, SLP: Slope

지 않다. 특히 지형의 기복이 심한 지역에서는 지형적 특성이 호우발생에 중요한 영향을 미친다. 많은 연구들에서 강우관측지점의 표고와 강우량 간의 선형관계를 다루고 있으나 이러한 선형기법은 소규모 기복의 지형단면이 거의 존재하지 않는 대규모 기복을 가지는 지역에서만 적절하며, 다른 조건을 가지는 지역에서는 표고뿐만 아니라 추가적인 지형특성이 유용할 수 있다. 이러한 지형특성변수로는 위도, 경도, 표고, 경사, 노출, 향, 방위각, 주곡률반경, 수역으로부터의 거리, 장벽효과 등이 있으며, 강우량과 이러한 지형특성변수의 단일관계는 특별한 경우를 제외하고는 일반적으로 상관성이 낮기 때문에 대부분의 연구들에서 중회귀분석을 적용하고 있다 (Kieffer and Bois, 2000; Bostan and Akyürek, 2007; Krähenmann and Ahrens, 2010).

따라서 본 연구에서는 확률강우량과 지형특성인자들 간의 관계를 규명하기 위하여 중회귀분석을 적용하였으며, 지형특성인자로서 위도 (Y), 경도 (X), 표고 (EL), 향 (ASP), 음영기복 (HLS) 및 경사 (SLP)를 고려하여 회귀모델의 설명변수를 설정하였다. 중회귀분석을 위해 사용된 선형모델은 Eq. (7)과 같다.

$$R_{100} = a_0 + a_1X + a_2Y + a_3EL + a_4ASP + a_5HLS + a_6SLP \quad (7)$$

여기서, R_{100} 은 목적변수인 지속시간 24시간 100년 빈도 확률강우량, a_0 는 상수, $a_1 \sim a_6$ 는 회귀계수이다.

앞서 분석한 지형인자들 간의 상관계수에서 X~EL, EL

~HLS, EL~SLP가 각각 0.67, -0.60, 0.74로서 종속성을 보이는 것으로 분석되었으므로 중회귀분석에서 독립변수들 간의 종속성으로 인한 다중공선성 (multicollinearity) 때문에 회귀모델의 매개변수 추정 시 그 결과가 왜곡될 가능성이 있다. 본 연구에서는 지형인자들 간의 다중공선성을 진단하기 위하여 Eqs. (8) and (9)와 같은 CI (condition index)와 VIF (variance inflation factor)를 산정하였다. 일반적으로 VIF가 10 이상, CI가 30 이상이면 다중공선성 문제가 발생한다고 알려져 있기 때문에 Table 5로부터 다중공선성 문제가 발생함을 확인할 수 있다.

$$CI = \sqrt{\frac{\lambda_{\max}}{\lambda_{\min}}} \quad (8)$$

$$VIF_k = \frac{1}{1 - R_k^2}, \quad k = 1, 2, \dots, p-1 \quad (9)$$

여기서, λ_{\max} 와 λ_{\min} 은 각각 행렬 $\mathbf{X}^T\mathbf{X}$ 에 대한 고유치의 최대값과 최소값, \mathbf{X} 는 설명변수 행렬, R_k^2 은 종속변수가 \mathbf{X}_k 이고 \mathbf{X}_k 를 제외한 나머지 설명변수를 설명변수로 하는 회귀모델에 대한 결정계수, p 는 설명변수의 개수이다.

중회귀분석시 설명변수간의 다중공선성 문제가 발생할 경우 회귀계수에 대한 OLS (ordinary least square) 추정치가 불안정한 값을 가질 수 있기 때문에 본 연구에서는 회귀계수를 추정하기 위하여 OLS 대신 LASSO (least absolute shrinkage and selection operator)를 적용하였

Table 5. Condition Index and Variance Inflation Factor

FACTORS	X	Y	EL	ASP	HLS	SLP
VIF	2.48	1.31	13.93	1.20	4.06	6.59
CI	27,961,649					

Table 6. Regression Coefficients and Adjusted Coefficients of Determination for Linear Models

No.	MODELS	a ₀	a ₁ (X)	a ₂ (Y)	a ₃ (EL)	a ₄ (ASP)	a ₅ (HLS)	a ₆ (SLP)	R ^{*2}
1	R100~ASP	3.122E+02	-	-	-	8.655E-02	-	-	-0.001
2	R100~X+ASP	3.122E+02	-2.836E-05	-	-	9.061E-02	-	-	-0.065
3	R100~X+EL+ASP	3.122E+02	-4.688E-05	-	1.457E-03	9.166E-02	-	-	-0.139
4	R100~X+EL+ASP+SLP	3.122E+02	-3.980E-04	-	5.325E-02	1.185E-01	-	-4.418E-01	-0.148
5	R100~X+EL+ASP+HLS	3.122E+02	-1.214E-03	-	2.455E-01	1.401E-01	1.138E+00	-1.961E+00	-0.046
6	R100~X+Y+EL+ASP+HLS+SLP	3.122E+02	-2.049E-03	2.375E-04	4.598E-01	1.570E-01	2.381E+00	-3.694E+00	-0.059

* $R^{*2} = 1 - \frac{n-1}{n-p-1}(1-R^2)$, where R^2 : coefficient of determination, n : sample size, p : number of explanatory variables

Table 7. Summary Statistics of KED Predictions

(Unit : mm)

No.	CASES	MIN.	1st QUAN.	MEDIAN	MEAN	3rd QUAN.	MAX.
1	R100~EL	274.9	302.9	311.1	315.8	321.3	421.8
2	R100~ASP	247.6	291.6	304.4	307.3	318.4	423.9
3	R100~HLS	257.0	301.8	311.6	315.9	324.9	449.8
4	R100~SLP	270.2	305.9	312.7	317.6	322.8	430.3
5	R100~X	274.1	301.7	310.3	313.2	315.7	420.6
6	R100~Y	271.7	302.5	312.7	314.5	318.7	420.6

다. LASSO는 다중공선성 문제를 완화시킬 수 있는 대표적인 축소기법 (shrinkage method)으로서 회귀계수에 패널티를 부여하여 그 크기를 축소시키고 잔차의 제곱합을 최소화시키며, 동시에 변수선택까지 가능한 기법이다. LASSO 기법에 대한 자세한 내용은 Tibshirani (1996)를 참고할 수 있다.

Table 6은 중회귀분석 결과로서 회귀모델에 대한 조정결정계수 R^{*2} 의 값이 모두 음의 값을 가지고 있음을 알 수 있다. 조정결정계수는 모델이 반응을 예측하는데 도움이 되지 않는 항을 포함할 경우 음의 값을 가질 수 있으며, 그 값이 1에 가까울수록 우수한 모델적합을 나타내기 때문에 본 연구에서 고려한 회귀모델의 경우 통계학적 효용성이 없는 것으로 분석되었다.

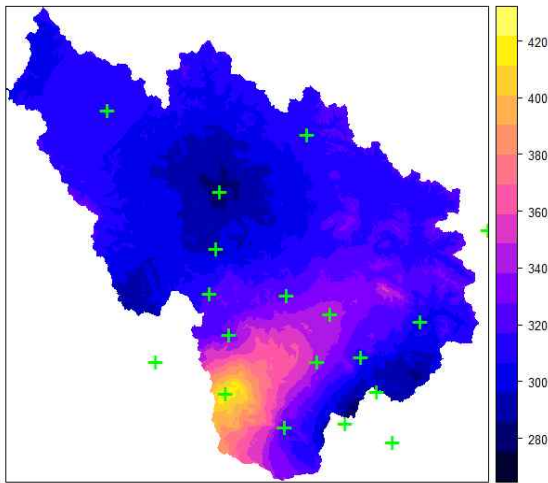
3.3 크리깅 분석

크리깅 분석에 있어서 베리오그램의 추정을 위하여 먼저 경험베리오그램을 작성하고 여러 이론베리오그램에 적합시켜 가장 적합한 이론베리오그램과 그 매개변수들을 추정하였다. 이론베리오그램으로서 지구통계학적 분석에서 널리 적용되고 있는 구형모델을 채택하였으며, 베

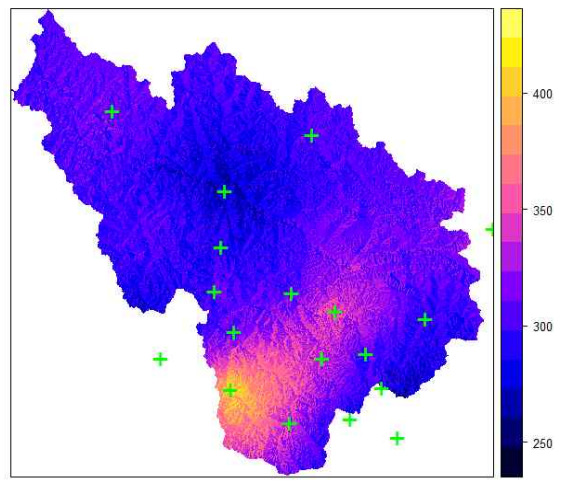
리오그램의 적합을 위해 가중최소제곱법을 적용하였다. 그리고 크리깅 기법으로서 단순크리깅, 보통크리깅, 일반크리깅 및 KED 기법을 적용하였다.

본 연구에서는 KED 기법의 적용을 위하여 지형특성인자를 보조변수로 고려하였으며, 앞서 분석된 중회귀분석 결과 조정결정계수 R^{*2} 의 값이 매우 낮아 회귀모델의 유효성이 없는 것으로 분석되었기 때문에 보조변수의 조합을 고려하지 않고 단일 보조변수만을 고려하여 KED 기법을 적용하였다.

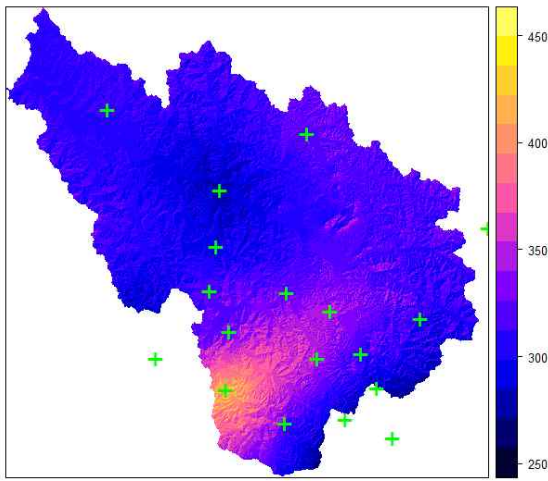
Table 7은 KED 기법을 적용하여 추정된 확률강우량의 공간분포에 대한 요약통계치를 정리한 것이다. 먼저 공간분포에 대한 평균값을 살펴보면 KED (R100~ASP)를 제외한 나머지 경우들은 대체로 비슷한 결과를 나타내는 것으로 분석되었다. 범위의 경우 KED (R100~HLS)가 192.8mm로서 가장 큰 범위를 나타냈으며, KED (R100~EL), KED (R100~X), KED (R100~Y)는 각각 146.9, 146.5, 148.9mm로서 대체로 비슷한 범위를 나타내는 것으로 분석되었다. 사분위수 범위는 KED (R100~ASP)가 26.8mm, KED (R100~HLS)가 23.1mm로서 다른 경우에 비해 다소 크게 나타나는 것으로 분석되었다. 한편, Fig. 3은 KED 기법



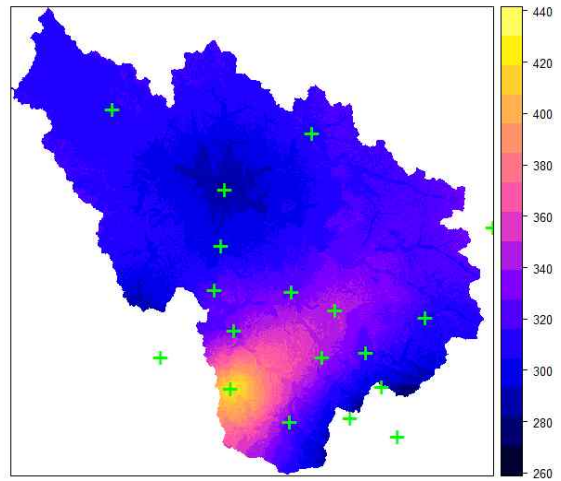
(a) KED(R100~EL)



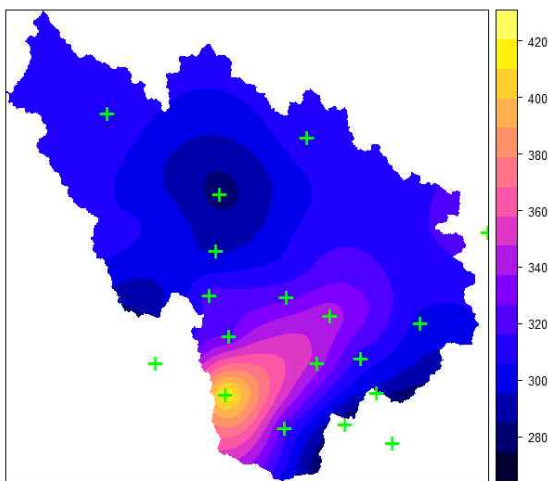
(b) KED(R100~ASP)



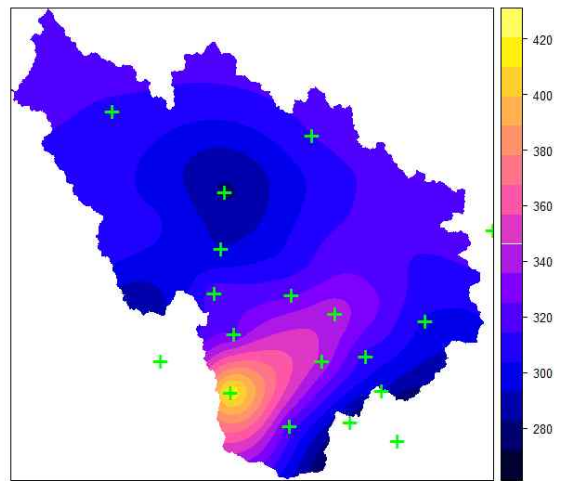
(c) KED(R100~HLS)



(d) KED(R100~SLP)



(e) KED(R100~X)



(f) KED(R100~Y)

Fig. 3. Prediction Maps for KED

Table 8. Cross Validation of KED and Simple and Ordinary Kriging

(Unit : mm)

NO.	METHODS		RMSE	ME
1	KED	R100~EL	36.407	-2.532
2		R100~ASP	38.707	-0.455
3		R100~HLS	39.178	1.072
4		R100~SLP	38.704	-0.547
5		R100~X	39.223	-1.096
6		R100~Y	37.767	-2.330
7	Simple Kriging		36.531	-2.836
8	Ordinary Kriging		37.277	-1.281
9	Universal Kriging		39.906	-2.316

Table 9. Comparison of Summary Statistics of Rainfall Spatial Distribution

(Unit : mm)

No.	METHODS		MIN.	1st QUAN.	MEDIAN	MEAN	3rd QUAN.	MAX.	
1	Deterministic	Arithmetic Mean	-	-	-	312.2	-	-	
2		Thiessen Polygon	-	-	-	314.5	-	-	
3		IDW (Weight=2.0)	278.5	307.6	311.6	314.2	318.0	420.8	
4		Trend Surface (Degree=1.0)	303.8	310.4	313.8	314.2	318.1	326.3	
5		Spline (Regularized, $\tau=0.1$)	190.5	292.2	312.6	313.6	336.0	434.8	
6	Geostatistical	Simple Kriging	274.7	302.9	312.1	314.4	317.6	420.6	
7		Ordinary Kriging	274.2	301.5	309.5	313.0	315.9	420.6	
8		Universal Kriging	271.6	302.5	312.5	314.4	318.7	420.6	
9		KED	R100~EL	274.9	302.9	311.1	315.8	321.3	421.8
10			R100~ASP	247.6	291.6	304.4	307.3	318.4	423.9
11			R100~HLS	257.0	301.8	311.6	315.9	324.9	449.8
12			R100~SLP	270.2	305.9	312.7	317.6	322.8	430.3
13			R100~X	274.1	301.7	310.3	313.2	315.7	420.6
14	R100~Y		271.7	302.5	312.7	314.5	318.7	420.6	

을 적용하여 추정된 확률강우량의 공간분포를 나타낸 것으로서 각 경우에 대하여 대체로 비슷한 공간패턴을 나타냄을 알 수 있다.

다음으로 본 연구에서는 KED 기법과 단순크리깅, 보통크리깅 및 일반크리깅에 대한 상세한 비교를 위하여 각 기법별로 Leave-one-out 교차검증을 수행하였다. Leave-one-out 교차검증은 K-fold 교차검증기법의 하나로서 전체 표본 N 개 중 1개의 자료만을 제외한 $N-1$ 개로 모델을 구축한 후 나머지 1개의 표본으로 검증하는 기법이다. Table 8은 교차검증에 대한 결과로서 RMSE (root mean square error)의 경우 KED (R100~EL)이 36.407 mm, 단순크리깅이 36.531 mm로서 다른 기법들에 비해 낮은 결과를 보이고 있으며, ME (mean error)의 경우 KED (R100~ASP)가 -0.455 mm, KED (R100~SLP)가 -0.547 mm로서 가장 낮은 결과를 보이고 있으나 대체로 RMSE 및

ME가 비슷한 것으로 분석되었다. 이로부터 KED 기법의 경우 비록 공간과정의 추세에 대해 다소 정도가 낮은 단순한 관계를 적용하였으나 단순크리깅, 보통크리깅 및 일반크리깅과 비슷한 정도로 공간예측을 수행하는 것으로 분석되었다.

Table 9는 강우공간분포 및 유역평균강우량 추정결과를 비교한 것으로서 KED 기법의 적용결과와 비교를 위해 확정론적 공간보간기법인 Thiessen 다각형, IDW, Trend Surface 및 Spline 기법, 그리고 지구통계학적 기법인 단순크리깅, 보통크리깅 및 일반크리깅을 적용하였다. Table 9로부터 Spline, KED (R100~ASP) 및 KED (R100~HLS)의 경우 확률강우량에 대한 공간분포의 범위 및 사분위 범위가 다른 기법들에 비해 다소 크게 나타났다. 또한, 경향면 분석의 경우 다른 기법들에 비해 범위 및 사분위 범위가 매우 작은 것으로 분석되었다. 그 밖의 기

법들은 대체로 비슷한 범위를 가지는 것으로 분석되었으며, 유역평균강우량은 KED (R100~ASP)의 경우 다른 기법들에 비해 다소 작은 반면 KED (R100~ASP)를 제외한 모든 기법들은 거의 비슷한 유역평균강우량을 나타내는 것으로 분석되었다.

비교결과 비교대상기법들은 공간분포에 대한 추정결과에서 범위의 차이가 다소 있기는 하지만 대체로 비슷한 결과를 나타내는 것을 알 수 있다. 유역평균강우량의 경우도 마찬가지로 거의 비슷한 결과를 나타내는 것을 알 수 있다. 그러나 비교대상기법들 중 확정론적 기법들은 지구통계학적 기법들과는 달리 강우관측치의 공간상관을 고려하지 않는 기법들이며, KED 기법을 제외한 나머지 기법들은 강우관측지점에서의 강우관측치만을 이용하는 기법들이기 때문에 유역전체에 대한 강우공간분포를 완전하게 나타낸다고는 할 수 없다. KED 기법의 경우 보조변수로서 유역의 지형특성인자들을 크리깅 시스템 내에 포함시키기 때문에 다른 기법들에 비해 상대적으로 유역전체 대한 강우공간분포를 보다 현실적으로 반영할 수 있다.

그러나 KED 적용에 있어서 대상 강우관측소의 수, 즉 표본수가 적고 앞서 분석한 바와 같이 강우측정지점에서의 지형특성인자들과 확률강우량간의 상관관계가 강하지 않기 때문에 KED 기법이 가지는 이론적인 이점을 실제 분석에 반영하는데 있어서 다소 한계가 있는 것으로 판단된다. 따라서 향후 지형특성인자에 대한 정확도 향상, 확률강우량과 지형특성인자간의 정밀한 통계학적 모델구축, 기타 특성인자들의 도입 등과 관련된 추가적인 연구가 필요할 것으로 판단되며, 본 연구의 경우 특정지역에 국한된 결과이므로 KED 기법과 관련한 일반적인 결과를 얻기 위하여 다양한 유역을 대상으로 한 비교연구가 또한 수행되어야 할 것으로 판단된다.

4. 결론

본 연구에서는 강우공간분포 및 유역평균강우량 추정시 신뢰도를 향상시키는데 있어서 보조변수 사용의 유효성을 평가하기 위하여 위천유역을 대상으로 KED 기법을 적용하였으며, 그 결과를 확정론적 공간보간기법인 Thiessen 다각형, IDW, 경향면 분석 및 Spline과 지구통계학적 공간보간기법인 단순크리깅, 보통크리깅 및 일반크리깅과 비교하였다.

확률강우량과 단일 보조변수로서 지형특성인자들과의 관계를 고려한 KED 기법의 적용 결과 강우공간분포 및 유역평균강우량의 추정에 있어서 확정론적 공간보간기법 및 크리깅 기법과 대체로 비슷한 결과를 나타내는 것으로 분석되었으며, KED 기법과 크리깅 기법에 대한 교차검증

결과 KED (R100~EL)이 가장 좋은 결과를 나타내고는 있으나 다른 기법들과 비교했을 때 큰 차이를 보이지 않는 것으로 분석되었다.

KED 기법은 보조변수로 사용된 인자들의 정확도, 강우량과 특성인자들 간의 상관관계 등에 영향을 받기 때문에 향후 관련 추가연구를 통해 보완되어야 할 것이며, 본 연구결과는 특정지역에 국한된 결과이므로 일반적인 결과를 얻기 위하여 다양한 유역을 대상으로 한 비교연구 또한 수행되어야 할 것으로 판단된다.

감사의 글

본 연구는 소방방재청에서 주관하고 있는 자연재해저감기술개발사업단 연구개발사업의 일환으로 수행된 “지구단위홍수재해 정보시스템 및 침수재해저감기술개발”의 일부로서 본 연구의 재정지원에 감사드립니다.

참고문헌

- Ahmed, S., and de Marsily, G. (1987). “Comparison of geostatistical methods for estimating transmissivity using data on transmissivity and specific capacity.” *Water Resources Research*, Vol. 23, No. 9, pp. 1717-1737.
- Bostan, P.A., and Akyürek, Z. (2007). “Exploring the mean annual precipitation and temperature values over Turkey by using environmental variables.” *ISPRS Joint Workshop of Visualization and Exploration of Geospatial Data*, University of Applied Sciences, Stuttgart, Germany.
- Bourennane, H., King, D., and Couturier, A. (2000). “Comparison of kriging with external drift and simple linear regression for predicting soil horizon thickness with different sample densities.” *Geoderma*, Vol. 97, No. 3-4, pp. 255-271.
- Buytaert, W., Celleri, R., Willems, P., Bièvre, B.D., and Wyseure, G. (2006). “Spatial and temporal rainfall variability in mountainous areas: A case study from the South Ecuadorian Andes.” *Journal of Hydrology*, Vol. 329, No. 3-4, pp. 413-421.
- Cole, S.J., and Moore, R.J. (2008). “Hydrological modeling using raingauge- and radar-based estimators of areal rainfall.” *Journal of Hydrology*, Vol. 358, No. 3, pp. 159-181.
- Daly, C. (2002). *Variable influence of terrain on pre-*

- precipitation patterns: Delineation and use of effective terrain height in PRISM.* Oregon State University, Corvallis, available at: <http://www.prism.oregonstate.edu/pub/prism/docs/effectiveterrain-daly.pdf>.
- Hartkamp, A.D., de Beurs, K., Stein, A., and White, J.W. (1999). *Interpolation techniques for climate variables.* 99-01, Wageningen Agricultural University, Wageningen.
- Haylock, M.R., Hofstra, N., Klein Tank, A.M.G., Klok, E.J., Jones, P.D., and New, M. (2008), "A European daily high-resolution gridded dataset of surface temperature and precipitation." *Journal of Geophysical Research*, Vol. 113, D20119.
- Kieffer, W.A., and Bois, P. (2000). "Topographic effects on statistical characteristics of heavy rainfall and mapping in the French Alps." *Journal of Applied Meteorology*, Vol. 40, pp. 720-740.
- Krähenmann, S., and Ahrens, B. (2010). "On daily interpolation of precipitation backed with secondary information." *Advances in Science and Research*, Vol. 4, pp. 29-35.
- Matheron, G. (1969). *Le Krigeage Universel.* Vol. 1, Cahiers du Centre de Morphologie Mathématique, École des Mines de Paris, Fontainebleau, p. NA.
- Moulin, L., Gaume, E., and Obled, C. (2009). "Uncertainties on mean areal precipitation: Assessment and impact on streamflow simulations." *Hydrology and Earth System Sciences*, Vol. 13, pp. 99-114.
- Odeh, I.O.A., McBratney, A.B., and Chittleborough, D.J. (1995). "Further results on prediction of soil properties from terrain attributes: Heterotopic cokriging and regression-kriging." *Geoderma*, Vol. 67, No. 3-4, pp. 215-226.
- Tibshirani, R. (1996). "Regression Shrinkage and Selection via the Lasso." *Journal of the Royal Statistical Society, Series B (Methodological)*, Vol. 58, No. 1, pp. 267-288.
- Wackernagel, H. (2003). *Multivariate geostatistics.* 3rd Edition, Springer-Verlag.
- Webster, R., and Oliver, M.A. (2001). *Geostatistics for environmental scientists, statistics in practice.* Wiley, Chichester, p. 265.
- Yatagai, A., Arakawa, O., Kamaguchi, K., Kawamoto, H., Nodzu, M.I., and Hamada, A. (2009). "A 44-year daily gridded precipitation dataset for Asia based on a dense network of rain gauges." *SOLA: Scientific Online Letters of the Atmosphere*, Vol. 5, pp. 137-140.

논문번호: 10-052	접수: 2010.06.30
수정일자: 2010.08.05/08.12	심사완료: 2010.08.12